Traditional Festival Cultural Tagging System and Tacit Knowledge Mining Post-Print

Authors: Su Ruizhu, Yan Jingya, Ouyang Jian

Date: 2023-04-01T16:15:48+00:00

Abstract

[Objective/Significance] The complexity inherent in traditional festival cultural expression often hinders the excavation of deeper connotations. This study employs knowledge organization methods to construct a tagging system for organizing and representing traditional festival cultural knowledge, thereby mining tacit knowledge underlying festival culture. [Method/Process] First, the structure and attributes of the traditional festival cultural tagging system are designed and defined. Second, web data is crawled and organized. Finally, underlying tacit associations are excavated from multiple dimensions. [Results/Conclusion] A tacit knowledge mining method based on the traditional festival cultural tagging system is proposed. Experimental results align with objective patterns, demonstrating its effectiveness and reliability.

Full Text

Traditional Festival Culture Tag System and Tacit Knowledge Mining

Su Ruizhu[1], **Yan Jingya**[1], **Ouyang Jian**[2] [1] School of Management, Guangxi University for Nationalities, Nanning 530006 [2] Library of Shanghai International Studies University, Shanghai 201620

Abstract: [**Purpose/Significance**] The complexity of traditional festival culture expression often makes it difficult to extract deeper meanings. This study utilizes knowledge organization methods to construct a tag system for organizing and expressing traditional festival culture knowledge, thereby mining the tacit knowledge behind festival culture. [**Method/Process**] First, the structure and attributes of the traditional festival culture tag system were designed and defined. Second, network data was crawled and collated. Finally, tacit relationships behind the data were excavated from multiple dimensions. [**Result/Conclusion**] This paper proposes a tacit knowledge mining method based

on a traditional festival culture tag system. Experimental results are consistent with objective laws, proving its validity and reliability.

Keywords: traditional festival; tag system; tacit knowledge; knowledge mining

---

Festival culture represents a nation's important cultural heritage, rich in historical and cultural connotations, and collectively showcases people's spiritual beliefs, aesthetic tastes, ethical relationships, and consumption habits [1]. UNESCO's Convention for the Safeguarding of the Intangible Cultural Heritage includes festive activities within its scope of protection [2]. Traditional festival culture constitutes a vital component of China's folk culture. As a ritual connecting emotions and inheriting culture in daily life, promoting festival culture is an important way to enhance national cohesion, a prerequisite and foundation for developing national culture, and a powerful force for creating a harmonious society. Therefore, researching and mining traditional Chinese festivals is of great significance for better revealing the cultural meaning and spiritual connotations of festivals. This paper uses theories and methods of knowledge organization from information science to conduct knowledge mining of Chinese traditional festivals, representing a beneficial exploration of festival culture gene mining from the perspective of knowledge organization.

1. Current State of Festival Culture Mining

Traditional culture plays a positive leading role in a country's overall development. In recent years, as China has increased its efforts to protect traditional culture, traditional festival culture has become closely related to public life. Cultural content changes with time, region, and population, and the environment for protecting, organizing, and utilizing folk culture has improved. Not only have rich ethnic cultural resource databases been established, promoting the development of traditional cultural undertakings, but they have also laid the foundation for mining traditional cultural values. Traditional cultural knowledge mining has gradually attracted scholars' attention. The deep knowledge of traditional festivals involves multi-angle knowledge extraction of festival time, population, region, origin, etc., with each angle involving deeper content such as history, geography, and humanities. Moreover, festival culture exists more in public consciousness and expression than in specialized cultural contributions and records. Therefore, compared with other types of intangible cultural knowledge mining, traditional festival culture knowledge mining has greater semantic complexity.

Given the dispersed and diversified nature of traditional festival culture knowledge, some scholars have used various methods to describe and represent the complex knowledge relationships in traditional festival culture. For example, some scholars propose that existing ontological frameworks can be applied to reveal tacit knowledge in traditional festival culture. In existing research, scholars such as Hao Tinglei [3], Wei Hao et al. [4], and Ma Xuming et al. [5] have pro-

posed building ontologies for traditional festivals; Ren Tingyan et al. suggested using XML clustering to mine ethnic festival culture [6]. However, research on mining and revealing tacit knowledge in traditional festival culture through ontology has two shortcomings: First, although it may essentially reveal and deeply mine the tacit knowledge behind traditional festivals, it requires a large number of relevant researchers to first build an ontology knowledge base around the knowledge space, then determine knowledge units for different themes, resulting in high implementation difficulty and cost [7]. Second, attribute values in traditional festival ontologies are difficult to update in a timely manner as traditional festival content changes with time, region, and population.

Therefore, this paper applies knowledge organization theories and methods to construct a traditional festival culture tag system and mine the relevance of labeled festival culture, deeply revealing the tacit knowledge hidden in traditional festival culture resources. First, knowledge organization methods are used to build a tag system to express traditional festival culture; second, network data is collected and extracted; finally, with support from certain data, tacit knowledge behind it is mined from multiple dimensions. This method can not only make full use of traditional festival culture tags on the internet, overcoming to some extent the high labor costs required to build ontology knowledge bases, but also alleviating the difficulties that traditional festival culture development brings to research.

2. Structure and Attributes of Traditional Festival Culture Tag System

2.1 Structure and Connotation of Traditional Festival Culture Tag System

In this paper, "traditional festivals" refer to festivals with long histories that can reflect the characteristics of Chinese culture, such as Spring Festival and Dragon Boat Festival, excluding various modern commercial tourism culture festivals such as Qingdao Golden Beach Cultural Tourism Festival.

Due to the fragmented nature of festival culture content, much of which is hidden in various classics, articles, or myths and legends, making it difficult to extract implicit relationships. Therefore, this paper first defines multiple attributes for each traditional festival, such as festival date, festival activities, festival diet, etc., to comprehensively reflect festival culture. Second, festival culture content is refined and expressed in tag form, which helps display festival characteristics, facilitates discovering relationships between attributes and tags, and thereby mines the tacit knowledge behind traditional festival culture, as shown in Figure 1 [FIGURE:1].

The traditional festival culture tag system is defined as S, where S = (FS, TS, OS). FS = {fS_1, fS_2, fS_3, …, fS | n [1, N]} is the set of various traditional festivals, with N being the total number of traditional festivals; TS = {tS_1,

tS $_2$, tS $_3$, …, tS | n [1, N], x [1, X]} is the set of all tags in the traditional festival culture tag system, which is the core content of this system, with X being the number of tags corresponding to the nth traditional festival culture; OS = {os$_1$, os$_2$, os$_3$, …, os | m [1, M]} is the set of data sources for traditional festival culture tags, where M is the number of data sources.

In the traditional festival culture tag system, if two traditional festivals are simultaneously revealed by multiple identical tags, this indicates that there is an important internal relationship between these two traditional festivals, or that some of their attributes may have deeper connections. For example: Linwang Festival, Dragon Killing Festival, and Tea Song Festival are all traditional festivals of the Dong ethnic group, so these festivals share the same "Dong" tag, thus proving there is a certain relationship between them; for example, ethnic groups such as Zhuang, Li, Yi, and Yao celebrate their traditional festivals on the third day of the third lunar month. Although the festival names differ, the festival dates are the same, indicating a certain relationship between these festivals or ethnic groups. Using the same method, internal relationships between different festivals, different ethnic groups, and different religions can be identified in terms of dietary habits, primitive beliefs, production methods, historical origins, ethnic migration, and traditional culture. Moreover, the higher the frequency of identical tags, the greater the correlation between two traditional festivals or festival attributes. Therefore, based on the correlation degree of tags possessed by different festivals, the commonalities and differences between different ethnic and religious groups can be summarized and distinguished.

2.2 Core Attributes and Levels of Traditional Festivals

This paper refers to the seven-step method, one of the ontology engineering construction methods, to analyze traditional festival culture content and determine tag attributes. The seven-step method was developed by Stanford University for domain ontology construction. The seven steps are: Determine the domain and scope of the ontology; Consider reusing existing ontologies; List important terms in the ontology; Define classes and class hierarchy; Define class attributes; Define facets of attributes; Create instances [8]. According to actual needs, the analysis of traditional festival tag attributes is mainly divided into the following four steps:

(1) Determine the domain and scope of traditional festival culture. To build a tag system, the domain and scope expressed by the tag system must first be determined. The domain of traditional festival culture refers to the special descriptions, behaviors, and populations involved in festivals during their emergence and evolution that distinguish them from other dates. One purpose of building a traditional festival culture tag system is to clearly express traditional festival culture, so the scope covered by traditional festival culture must be determined and the content to be expressed clarified to stabilize the knowledge expression structure.

(2) Determine the scope of core festival attributes. Festival attributes are general terms for certain characteristics of festivals and represent different perspectives for studying and viewing festivals, while festival culture is the collection of these attributes. Through summarization of the above materials, core attributes are proposed: "Festival Name," "Festival Alias," "Festival Date," "Festival Activities," "Festival Diet," "Related Figures," "Festival Ethnic Group," "Festival Origin," "Origin Location," "Origin Time," "Myths and Legends," "Popular Regions," and "Related Religion."

(3) Clarify core festival attributes. Festival Name: A name given by official institutions or most widely circulated for a festival. Festival Alias: Names other than "Festival Name" for a festival. Festival Date: The date or time period when the festival occurs. Festival Activities: Behaviors and activities used by the public for celebration and commemoration during the festival. Festival Diet: Special foods consumed by the public during the festival. Related Figures: Real historical figures related to the festival. Festival Ethnic Group: Ethnic groups that celebrate the festival. Festival Origin: The reason for the festival's emergence, such as events commemorated, tasks, or belief cultures about the festival's origin. Origin Location: The region where the festival originated. Origin Time: The era or year when the festival originated. Myths and Legends: Non-historical stories related to the festival's emergence and evolution. Popular Regions: Areas where the festival is currently popular. Related Religion: Religions related to the festival's origin or evolution.

(4) Hierarchical classification of core festival attributes. Since some tags exist on the internet in the form of simple words or phrases that are convenient for data extraction, while others need to be extracted from article content and are difficult to extract automatically, often requiring manual collation, this paper classifies attributes into three levels according to extraction difficulty: shallow attributes, middle attributes, and deep attributes (see Table 1).

Shallow attributes refer to attributes that are already labeled in data sources and can directly form tags using raw data; middle attributes refer to attributes that require extracting words or phrases from article paragraphs to form tags; deep attributes refer to attributes that require understanding and summarizing article or story content to form tags, which are greatly influenced by data processing personnel's subjectivity. This paper selects data from three data sources to form tags based on 13 attributes, constructing the traditional festival culture tag system model S'.

3. Construction of Traditional Festival Culture Tag System Model

The research direction of using tag systems for knowledge mining originated from the concept of "Folksonomy." Folksonomy means folk classification, al-

lowing the public to classify and label information resources. Based on this, concepts such as “social tags” and “social bookmarks” emerged. Social tags are a research hotspot in China’s library and information science field, widely applied in information retrieval, data visualization, personalized services, knowledge mining, and other application fields. Many scholars in the library and information field have studied tag information from perspectives such as complex networks, ontologies, and clustering, exploring relationships and deep semantic information between tags [9]. They found that tags have various relationships, such as co-occurrence relationships, hierarchical relationships, and correlation relationships. Strengthening the mining of relationships between tags on a certain theme helps reveal tacit knowledge of that theme. Therefore, this paper proposes tagging traditional festival culture and revealing relationships between tags by constructing a traditional festival culture tag system, providing ideas for traditional festival culture research.

3.1 Construction Process

The first step in building a traditional festival culture tag system is attribute definition, which defines the core concepts of traditional festival culture to form the upper structure of festival culture tags. Simultaneously, data sources are selected according to research purposes. The second step is to extract relevant data from data sources, summarize and integrate it, and form tags from the collated data corresponding to the upper attributes. Thereby, the traditional festival culture tag system model S’ can be established, as shown in Figure 2

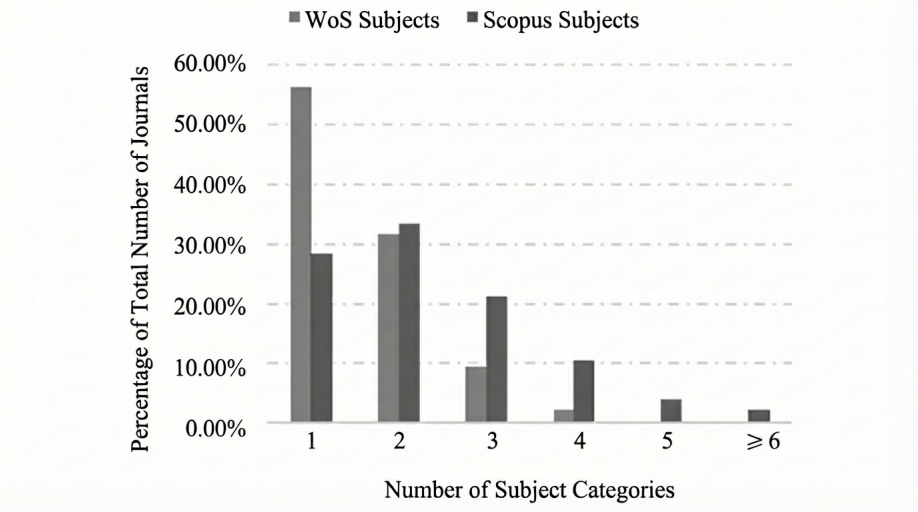

Figure 1: Figure 2

.

3.2 Data Sources and Extraction

3.2.1 Data Sources Traditional festival culture exists in official authoritative databases and in sources like Baidu Baike that can be edited by the public. Official authoritative databases are published or built by authoritative institutions in the traditional festival culture field, with scientific, accurate, and highly credible attributes, but with smaller data volumes; encyclopedia websites have larger data volumes and certain credibility.

(1) Authoritative Databases. "China Civilization Network" is hosted by the Publicity Department of the CPC Central Committee and the Central Civilization Committee [10]; "Chinese Festival Network" is guided by the State Ethnic Affairs Commission and hosted by the Ethnic Festivals Professional Committee of the Chinese Anthropological and Ethnological Research Association [11]. Both websites are managed by official institutions, making their data highly authoritative and authentic.

(2) Baidu Baike. "Baidu Baike" is an online encyclopedia platform containing massive amounts of high-quality semi-structured data, including article titles, category tags, information boxes, and other standardized knowledge. Most entries are edited by humans, and edited content can only be published after review, giving it certain credibility.

3.2.2 Data Extraction This paper adopts two data extraction strategies: Strategy one, for web pages with relatively clear structure where network tags can be directly obtained, crawler software is used to identify and extract traditional festival culture attributes and attribute values. This paper develops different extraction templates and rules for different data websites, using the GooSeeker crawler software to achieve automatic crawling of web page information for rapid and effective extraction of candidate data from different data source websites. First, a crawling rule needs to be established for sample pages (e.g., "China Civilization Network - January Festivals" page content), which maps tags in the crawling rule to web page information by labeling the fields to be collected on sample pages, then adds crawling clues to crawl information on other pages with the same structure (e.g., "China Civilization Network - February Festivals" page content). Second, since Baidu Baike festival entry URLs are difficult to obtain directly from web pages, keyword lists are needed to construct clues and crawl relevant page information. This paper obtained and collated 263 traditional festivals from China Civilization Network and Chinese Festival Network to form a keyword directory. Strategy two, for attributes and attribute values that crawler software cannot automatically identify, this paper adopts manual extraction methods, manually extracting traditional festival attribute values from article paragraphs and corresponding them with attributes to form descriptions of traditional festivals.

3.3 Tag Integration

Since data from the three data sources are all semi-structured, and there are differences in data structure and language expression among the three sources, with many cases of synonyms where different words express the same meaning (for example: "Name" and "Chinese Name" for festivals have different expressions but the same meaning), cleaning, deduplication, and other work is needed to reduce redundant knowledge and achieve data integration.

Attribute alignment is an important part of tag integration. Attribute alignment refers to the unification of attribute vocabulary in heterogeneous data sources. Since the data sources selected in this paper are from three different websites, different websites use different descriptions for the same attribute of the same entity in some attributes, but each data source website uses fixed descriptions for fixed attributes. Therefore, the correctness of attribute alignment among different data sources can be ensured by manually constructing a synonym mapping table [12]. For example, for the description of festival activities, Baidu Baike uses "Festival Activities," while China Civilization Network uses "Customs," so a unified attribute synonym mapping table {Festival Activities, Celebration Activities, Customs} can be established for these attributes, allowing tags under these attributes to uniformly correspond to the attribute "Festival Activities."

After attribute alignment is completed, tags will correspond one-to-one with attributes, filling tags into various attributes to form cultural tags for a single festival (see Figure 4

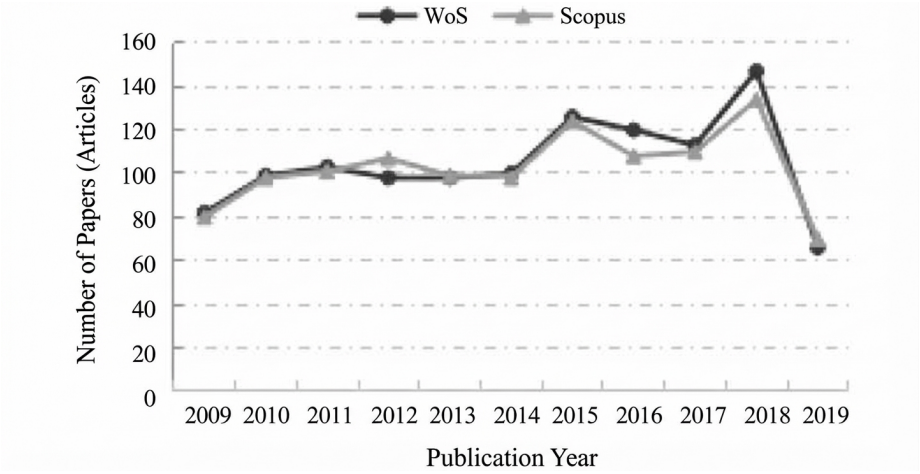

Figure 2: Figure 4

), using Spring Festival as an example:

Attribute: Festival Name; Tag: Spring Festival; Attribute: Festival Alias; Tags:

New Year, New Year's Day…; Attribute: Festival Date; Tag: First day of first lunar month; Attribute: Festival Ethnic Groups; Tags: Han, Manchu…; Attribute: Festival Activities; Tags: Pasting red couplets, New Year visits, staying up late…; Attribute: Festival Diet; Tags: New Year cake, spring rolls, dumplings…;

The partial content of the traditional festival culture tag system model S' constructed in this paper is shown in Table 2 . S' = (F'S, T'S, O'S), where F'S is the set of all traditional festivals in this paper, with 263 traditional festivals obtained and collated, N = 263: Spring Festival, etc.; T'S is the set of all tags in the traditional festival culture tag system in this paper, belonging to 13 attributes: "Festival Name," "Festival Alias," "Festival Date," "Festival Ethnic Group," "Festival Activities," "Festival Diet," "Related Figures," "Popular Regions," "Festival Origin," "Origin Location," "Origin Time," "Myths and Legends," and "Related Religion"; O'S is the set of data sources for traditional festival culture tags, M = 3: "China Civilization Network - Traditional Festivals," "Chinese Festival Network - Ethnic Festivals," and Baidu Baike.

4. Tacit Knowledge Mining at Attribute Levels of Traditional Festival Culture

Based on the traditional festival culture tag system model S', this paper conducts statistical analysis of tag data and carries out in-depth mining of tacit knowledge at attribute levels of traditional festival culture. When mining tacit knowledge in traditional festival culture, tag data processing methods differ greatly across different attribute levels. Tags of shallow attributes have strong objectivity and stability and can directly undergo statistical analysis based on raw data, while tags of middle or deep attributes require further processing according to research needs. Since this only involves research on mining ideas, experiments are conducted on the traditional festival culture tag system from only four dimensions—festival date, festival ethnic group, festival diet, and festival origin—selected from the three attribute levels, with the experimental process and results briefly elaborated.

4.1 Tacit Knowledge Mining of Shallow Attributes

4.1.1 Tacit Knowledge Mining of Festival Date Dimension Festivals emerge because something or some change that needs to be remembered occurred on that day or around that day, so people give that date a different meaning from ordinary dates as a commemoration and reminder in daily life, usually cycling once a year. The calendar is one of humanity's most important inventions. People live in the cycle of seasons and natural changes, so they invented calendars to conveniently record the passage of time and changes of seasons, while festivals record the most important nodes in the river of time. Therefore, statistical analysis of festival date distribution throughout the year can clearly observe the most important time periods in people's lives, find relevant patterns, and explore underlying reasons. In the knowledge mining process,

summarization is the simplest mining method, and tacit knowledge mining of shallow attributes mainly uses data summarization methods.

After removing traditional festivals with zero tags in the festival date attribute or dates not calculated by the lunar calendar, 161 traditional festivals were obtained. The number of festivals in each month was calculated and imported into Excel to form a bar chart of monthly festival distribution (see Figure 5

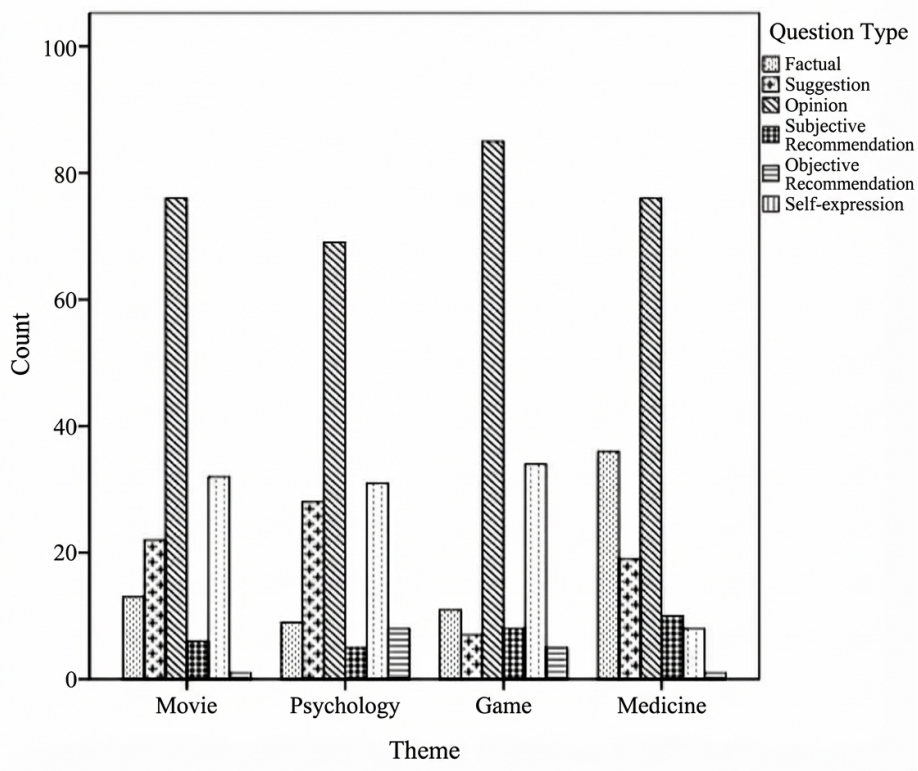

Figure 3: Figure 5

). Traditional festivals are mostly concentrated in the first half of the year, with fewer festivals in the second half. The first lunar month has the most traditional festivals, November has the fewest, and festivals in the twelfth lunar month are relatively numerous in the second half. The distribution pattern of festival dates is similar to the distribution pattern of agricultural solar terms. Since the first and twelfth lunar months are before and after the year-end transition, the cold weather in the twelfth month leaves agriculture idle, giving people time for various sacrifices and gathering activities; while celebrating the Spring Festival in the first month, people also begin planning for the new year's life and production, thus deriving many minor festivals, such as the Han traditional Little New Year (23rd day of twelfth lunar month) and Miao Jump Slope Festival (9th day

of first lunar month). As spring arrives and the weather warms, all things revive, and people begin busy agricultural days. The higher number of festivals in the first half of the year first serves to remind people of seasonal changes, such as Zhonghe Festival and Dong Planting Festival, which were established to urge people to farm in time and prepare for busy spring plowing. Second, farming was of extraordinary significance to people's survival in ancient times, and people often performed sacrificial rituals to pray for smooth farming, which over time produced festivals.

4.1.2 Tacit Knowledge Mining of Festival Ethnic Group Dimension
Festivals are important markers of ethnic identity. Festival culture contains cultural factors such as local customs, religious beliefs, and moral ethics in ethnic life, and is the long-term accumulation of an ethnic group's historical culture [13]. Different ethnic groups also celebrate the same festivals. If the phenomenon of "two ethnic groups celebrating the same festival" is defined as "there exists a festival relationship between the two ethnic groups," the more common festivals two ethnic groups share, the stronger the festival relationship between them. Therefore, by calculating the co-occurrence of ethnic tags, the strength of festival relationships between ethnic groups can be observed.

After removing data with zero or one ethnic tags from the raw data, the co-occurrence times of two ethnic groups in different festivals were calculated, and the festival ethnic group tags were analyzed and visualized using the social network analysis tool Gephi (see Figure 6

). In Figure 6, nodes represent ethnic tags, and connections between nodes indicate that this pair of ethnic groups co-occurs in one or more festivals. The more connections a node has, the more ethnic groups have festival relationships with that ethnic group, and the broader that ethnic group's festival relationship network; the thicker the line between two nodes, the stronger the festival relationship between that pair of ethnic groups. It can be seen that ethnic groups such as Manchu and Han have relatively broad festival relationship networks, while festival relationships between Dai and De'ang, Hui and Uzbek, etc. are relatively strong.

Due to geographical factors, ethnic groups living closer together often have stronger festival relationships. Observing Figure 6 reveals that the division of festival relationships among ethnic groups is quite similar to the north-south geographical division. Festival relationships are stronger among southern ethnic groups and among northern ethnic groups, while weaker between southern and northern ethnic groups. Moreover, the phenomenon that "festival relationships among northern ethnic groups are stronger than among southern ethnic groups" is discovered. Consulting relevant materials on the causes of differences between northern and southern ethnic groups reveals that southern China mostly has mountainous and hilly terrain, where rivers and terrain hinder interaction between ethnic groups. Second, northern ethnic groups in China mostly lived on nomadic hunting in early times, with frequent and extensive popula-

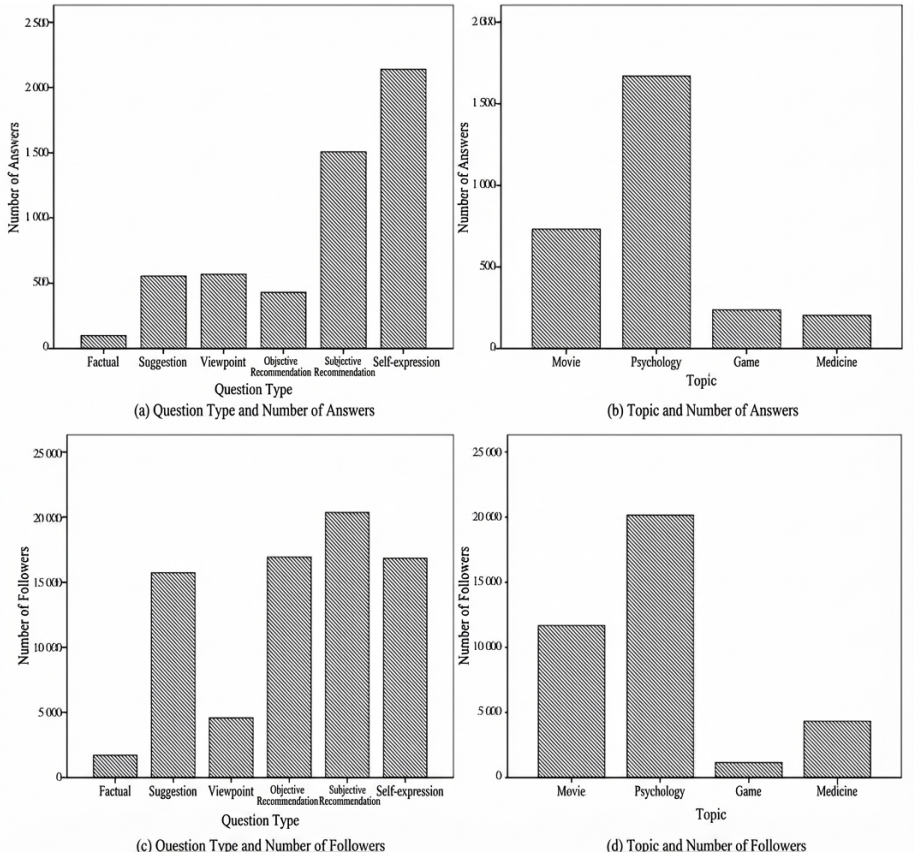

Figure 4: Figure 6

tion migration, while southern ethnic groups mostly had farming cultures with fixed residences. Finally, due to frequent warfare among northern ethnic minorities in early times, various tribes merged through conquest and integration, also causing current ethnic fusion. Therefore, although northern and southern ethnic groups have had considerable interaction throughout history, terrain, agricultural civilization, and warfare factors have also created certain north-south differences. Such distribution patterns of tags can effectively convey tacit knowledge behind traditional festival culture, and relevant researchers can conduct tacit knowledge mining work by analyzing traditional festival culture tags.

4.2 Tacit Knowledge Mining of Middle Attributes

Middle attribute tags are composed of words or phrases extracted from article paragraphs, with fixed content, but tags of different festivals' middle attributes often have various associations, which often imply tacit knowledge difficult to observe through tags themselves. Taking festival diet tags as an example, dietary culture is an important cultural representation of an ethnic group. An ethnic group's dietary habits are often related to its agricultural civilization and living region. Since people eat the most characteristic and meaningful foods of their ethnic group on festival days, festival diets can better reflect an ethnic group's cultural characteristics.

By statistically analyzing the distribution of festival diets among ethnic groups, the dietary characteristics of different ethnic groups and China's agricultural civilization can be discovered. Since some festivals (such as Spring Festival) are celebrated by multiple ethnic groups, festival diet tags of different ethnic groups in the same festival are first divided, and tags with unclear ethnic characteristics such as "fresh fruit tea snacks," "ancestor worship golden pig," and "tea snacks" are removed. Then some synonyms are merged (for example, "fried sanzi / fried sanzi" are all changed to "fried sanzi," and "new rice / new grain / new rice meal" are all changed to "new rice"), obtaining traditional festival diet tag sets for 56 ethnic groups. The co-occurrence of diet tags among ethnic groups is calculated and visualized using UCINET software (see Figure 7

). Points in Figure 7 represent diet tags, connections indicate co-occurrence of diet tag pairs, and more connections for a node in the network indicate a broader relationship network and higher centrality. Nodes with higher centrality in Figure 7 have larger areas. The color distinction clusters festival diet tags, and when the number of groups is 20, it is the optimal classification combination.

Observing Figure 7

, on the one hand, ethnic relationships can be seen from the clustering distribution patterns of diets: fried oil incense, fried sanzi, and other foods are beloved by ethnic groups such as Hui and Kazakh. These ethnic groups live in close regions and share the same religious culture. Yogurt, qeema, and other foods are more popular among ethnic groups such as Tibetan and Monba, while dairy products are also festival diets of Mongolian and other ethnic groups. These

Free Q&A:

$$Y_{\text{answer length}} = \beta_0 + \beta_1 X_{\text{question length}} + \beta_2 X_{\text{factual type}} + \beta_3 X_{\text{advisory type}} + \beta_4 X_{\text{subjective recommendation type}} + \beta_5 X_{\text{objective recommendation type}} + \beta_6 X_{\text{self-expression type}} + \beta_7 X_{\text{psychologys}} + \beta_7 X_{\text{psychology}} + \beta_8 X_{\text{games}} + \beta_9 X_{\text{medicine}} + \beta_{10} X_{\text{question posting time}} \quad \text{(Formula 1)}$$

$$Y_{\text{number of answers}} = \beta_0 + \beta_1 X_{\text{question length}} + \beta_2 X_{\text{factual type}} + \beta_3 X_{\text{advisory type}} + \beta_4 X_{\text{subjective recommendation type}} + \beta_5 X_{\text{objective recommendation type}} + \beta_6 X_{\text{self-expression type}} + \beta_7 X_{\text{psychology}} + \beta_8 X_{\text{games}} + \beta_9 X_{\text{medicine}} + \beta_{10} X_{\text{question posting time}} \quad \text{(Formula 2)}$$

$$Y_{\text{number of ratings}} = \beta_0 + \beta_1 X_{\text{question length}} + \beta_2 X_{\text{answer length}} + \beta_3 X_{\text{factual type}} + \beta_4 X_{\text{advisory type}} + \beta_5 X^{\text{subjective}}_{\text{recommendation type}} + \beta_6 X^{\text{objective}}_{\text{recommendation type}} + \beta_7 X_{\text{self-expression type}} + \beta_8 X_{\text{psychology}} + \beta_9 X_{\text{games}} + \beta_{10} X_{\text{medicine}} + \beta_{11} X_{\text{question posting time}} + \beta_{12} X_{\text{answer posting time}} \quad \text{(Formula 3)}$$

Paid Q&A:

$$Y_{\text{answer length}} = \beta_0 + \beta_1 X_{\text{number of historical answers}} + \beta_2 X_{\text{price}} + \beta_3 X_{\text{question length}} + \beta_4 X_{\text{factual type}} + \beta_5 X_{\text{advisory type}} + \beta_6 X^{\text{subjective}}_{\text{recommendation type}} + \beta_7 X^{\text{objective}}_{\text{recommendation type}} + \beta_8 X_{\text{self-expression type}} + \beta_9 X_{\text{psychology}} + \beta_{10} X_{\text{games}} + \beta_{11} X_{\text{medicine}} + \beta_{12} X_{\text{question posting time}} \quad \text{(Formula 4)}$$

$$Y_{\text{number of ratings}} = \beta_0 + \beta_1 X_{\text{number of historical answers}} + \text{price} + \beta_3 X_{\text{question length}} + \beta_4 X_{\text{answer length}} + \beta_5 X_{\text{factual type}} + \beta_6 X_{\text{advisory type}} + \beta_7 X^{\text{subjective}}_{\text{recommendation type}} + \beta_8 X^{\text{objective}}_{\text{recommendation type}} + \beta_9 X_{\text{self-expression type}} + \beta_{10} X_{\text{psychology}} + \beta_{11} X_{\text{games}} + \beta_{12} X_{\text{medicine}} + \beta_{13} X_{\text{question posting time}} \quad \text{(Formula 5)}$$

Figure 5: Figure 7

Free Q&A:

$$Y_{\text{answer length}} = \beta_0 + \beta_1 X_{\text{question length}} + \beta_2 X_{\text{factual type}} + \beta_3 X_{\text{advisory type}} + \beta_4 X_{\text{subjective recommendation type}} + \beta_5 X_{\text{objective recommendation type}} + \beta_6 X_{\text{self-expression type}} + \beta_7 X_{\text{psychologys}} + \beta_7 X_{\text{psychology}} + \beta_8 X_{\text{games}} + \beta_9 X_{\text{medicine}} + \beta_{10} X_{\text{question posting time}} \quad \text{(Formula 1)}$$

$$Y_{\text{number of answers}} = \beta_0 + \beta_1 X_{\text{question length}} + \beta_2 X_{\text{factual type}} + \beta_3 X_{\text{advisory type}} + \beta_4 X_{\text{subjective recommendation type}} + \beta_5 X_{\text{objective recommendation type}} + \beta_6 X_{\text{self-expression type}} + \beta_7 X_{\text{psychology}} + \beta_8 X_{\text{games}} + \beta_9 X_{\text{medicine}} + \beta_{10} X_{\text{question posting time}} \quad \text{(Formula 2)}$$

$$Y_{\text{number of ratings}} = \beta_0 + \beta_1 X_{\text{question length}} + \beta_2 X_{\text{answer length}} + \beta_3 X_{\text{factual type}} + \beta_4 X_{\text{advisory type}} + \beta_5 X^{\text{subjective}}_{\text{recommendation type}} + \beta_6 X^{\text{objective}}_{\text{recommendation type}} + \beta_7 X_{\text{self-expression type}} + \beta_8 X_{\text{psychology}} + \beta_9 X_{\text{games}} + \beta_{10} X_{\text{medicine}} + \beta_{11} X_{\text{question posting time}} + \beta_{12} X_{\text{answer posting time}} \quad \text{(Formula 3)}$$

Paid Q&A:

$$Y_{\text{answer length}} = \beta_0 + \beta_1 X_{\text{number of historical answers}} + \beta_2 X_{\text{price}} + \beta_3 X_{\text{question length}} + \beta_4 X_{\text{factual type}} + \beta_5 X_{\text{advisory type}} + \beta_6 X^{\text{subjective}}_{\text{recommendation type}} + \beta_7 X^{\text{objective}}_{\text{recommendation type}} + \beta_8 X_{\text{self-expression type}} + \beta_9 X_{\text{psychology}} + \beta_{10} X_{\text{games}} + \beta_{11} X_{\text{medicine}} + \beta_{12} X_{\text{question posting time}} \quad \text{(Formula 4)}$$

$$Y_{\text{number of ratings}} = \beta_0 + \beta_1 X_{\text{number of historical answers}} + \text{price} + \beta_3 X_{\text{question length}} + \beta_4 X_{\text{answer length}} + \beta_5 X_{\text{factual type}} + \beta_6 X_{\text{advisory type}} + \beta_7 X^{\text{subjective}}_{\text{recommendation type}} + \beta_8 X^{\text{objective}}_{\text{recommendation type}} + \beta_9 X_{\text{self-expression type}} + \beta_{10} X_{\text{psychology}} + \beta_{11} X_{\text{games}} + \beta_{12} X_{\text{medicine}} + \beta_{13} X_{\text{question posting time}} \quad \text{(Formula 5)}$$

Figure 6: Figure 7

ethnic groups either share the same cultural beliefs or have similar historical living environments, leading to similar dietary habits among them. On the other hand, from the perspective of centrality, foods such as zongzi and mooncakes have higher centrality, indicating that these foods are consumed by more ethnic groups. Zongzi and mooncakes are representative foods of China's two major traditional festivals—Dragon Boat Festival and Mid-Autumn Festival. Due to historical exchanges and integration among China's ethnic groups, the culture of celebrating the same festivals and tasting the same delicacies has emerged. Overall, since rice cultivation in China has a wide area and long history, and the status of rice cultivation in China's agricultural civilization determines the position of rice in the dietary culture of all ethnic groups in China, foods made from rice such as zongzi and glutinous rice cake are traditional representative foods of most ethnic groups in China. Although China is mainly agricultural, a small number of ethnic groups have unique nomadic cultures, such as Tibetan and Mongolian, and their unique nomadic culture is also reflected in their festival dietary culture.

4.3 Tacit Knowledge Mining of Deep Attributes

Deep attributes are summarized from article content and represent the basic feature expression and induction of festivals. Mining deep attribute tags can discover macro-level features of festivals. Taking festival origin as an example, China's traditional festivals are diverse and rich in content, constituting an important part of the Chinese nation's long historical culture. From ancient times to modern times, the development process of Chinese traditional festivals is also a long-term accumulation process of the Chinese nation's historical culture [14]. All of China's traditional festivals evolved from ancient or even prehistoric times. The formation of most ancient Chinese festivals was due to people's worship and belief in powerful forces in an era when science was not developed, such as ghost and deity legends, religious beliefs, natural solar terms, and heroic figures. Statistics on festival origins can reveal ancient people's primitive beliefs and ways of thinking.

Since many festival origins can no longer be found, and they often exist in classics, articles, and legends with long and complex content, making it difficult to mine tacit knowledge contained in festival origins, this paper synthesizes theories on festival origins and sources from scholars such as Mou Yuankui [15], Yuan Li et al. [16], and Wang Juan [17] to summarize festival origins into the following six major beliefs: ghost and deity belief, natural belief, religious belief, hero belief, ancestor belief, and multi-belief combination. By observing the proportion of various origins (see Figure 8

), it can be seen that no single belief accounts for more than half. Ghost and deity belief, hero belief, and natural belief account for relatively large proportions in ancient Chinese people's beliefs, while ancestor belief and religious belief account for smaller proportions, and multi-belief combination accounts for a relatively high proportion. Natural belief, ancestor belief, and ghost and deity

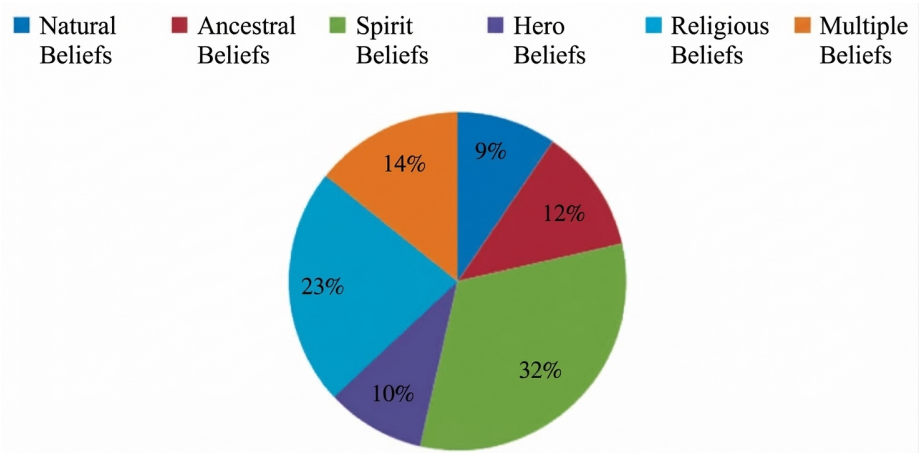

Figure 7: Figure 8

belief are deeply integrated. In the underdeveloped ancient era, when people could not scientifically explain some natural phenomena, they fabricated ghost and deity stories to rationalize these phenomena. Because people heavily depended on nature for production and life, the power of these ghosts and deities became very strong in people's eyes, and people had to take measures to please ghosts and deities for protection. Since festival origin tags are mostly legends or events about festival origins, they are mostly related to certain figures or deities and less related to ancestors. However, the tag "ancestor worship" in festival activities appears 16 times, thus revealing that although ancestor belief has a small relationship with festival origins, it has greatly influenced the development of festival culture. People often choose to worship ancestors on important days.

Mining tacit associations between ethnic groups contained behind traditional festival culture from a single attribute dimension can provide in-depth understanding of ethnic commonalities and differences from aspects such as belief culture and dietary culture. Mining content of a theme from multiple attribute dimensions simultaneously can achieve mutual support and verification. For example, observing the mining results of festival ethnic group tags and festival diet tags clearly shows that ethnic groups with large differences in residence also have large differences in diet. Single-dimensional mining is difficult to comprehensively display issues, making multi-dimensional mining work also necessary.

This paper proposes a tacit knowledge mining model for traditional festival culture tag systems from the perspective of knowledge organization. The process of tagging traditional festival culture is both an analysis of traditional festivals and knowledge organization and collation of cultural content. The traditional festival culture tag system can help relevant researchers describe and organize festival culture content in tag form, while the tag form also meets the needs of computer processing. It can not only make full use of traditional festival

culture tags on the internet, overcoming to some extent the high labor costs required to build ontology knowledge bases, but also alleviating the difficulties that traditional festival culture development brings to research. The tag system model from the perspective of knowledge organization can not only be used to research traditional festival culture but also has reference significance for research on China's traditional culture. The tag system provides a framework for more clearly organizing traditional culture, offers ideas for traditional culture expression, and helps scholars studying traditional culture more intuitively observe tacit associations behind traditional culture.

References

[1] Xiao Fang. Traditional Festivals: A Major National Cultural Heritage [J]. Journal of Beijing Normal University (Social Sciences Edition), 2005(5): 50-56. [2] International Department, Ministry of Culture. Compilation of Basic Documents of UNESCO's Convention for the Safeguarding of the Intangible Cultural Heritage [M]. Beijing: Foreign Languages Press, 2012: 7. [3] Hao Tinglei. Research on the Construction and Representation Method of Traditional Festival Knowledge Ontology [D]. Wuhan: Central China Normal University, 2011. [4] Wei Hao, Sun Yu. Research on the Construction of Ethnic Festival Domain Ontology Based on OWL [J]. Information Technology and Informatization, 2016(11): 87-89. [5] Ma Xuming, Wang Hairong. Research on the Construction Method of Hui Traditional Festival Ontology [J]. Electronic Technology and Software Engineering, 2018(5): 155-157. [6] Ren Tingyan, Luo Gang. Application of XML Clustering in Ethnic Festival Culture Mining [J]. Software Guide, 2015, 14(12): 140-141. [7] Cheng Xiufeng, Bi Chongwu, Li Chenglong. Research on Tacit Knowledge Association of Dance Intangible Cultural Heritage Based on SNA [J]. Library and Information Service, 2016, 60(2): 30-36. [8] Liu Renning, Li Yusheng. Domain Ontology Construction Method [J]. Journal of Wuhan Polytechnic University, 2008(1): 46-49, 53. [9] Li Xuhui, Li Yuanyuan, Ma Feicheng. Analysis of Main Issues in Social Tag Research in China's Library and Information Field [J]. Library and Information Service, 2018, 62(16): 120-131. [10] Editorial Department of China Civilization Network. China Civilization Network - About Us [EB/OL]. [2019-04-02]. http://www.wenming.cn/c/gywm/. [11] Ethnic Festivals Professional Committee of Chinese Anthropological and Ethnological Research Association. Chinese Festival Network [EB/OL]. [2019-04-02]. http://www.zhjqw.org/. [12] Wang Weiwei, Wang Zhigang, Pan Liangming, et al. Research on the Construction of Bilingual Film and Television Knowledge Graph [J]. Journal of Peking University (Natural Science Edition), 2016, 52(1): 25-34. [13] Qin Junying. Research on the Development and Utilization of Ethnic Minority Cultural Resources in "Cultural Life" Teaching [D]. Changsha: Hunan Normal University, 2014. [14] Luo Zerong. Research on Chinese Traditional Festival Culture from the Perspective of Historical Materialism [D]. Changsha: Hunan University, 2009. [15] Mou Yuankui. The Origin and Evolution of Chinese Seasonal Festivals [J]. Root Exploration, 1999(1): 9-12. [16] Yuan Li, Gu Jun. Chinese Folklore Tutorial [M]. Beijing: Guangming Daily

Press, 2003: 101-102. [17] Wang Juan. Introduction to Folklore [M]. Beijing: Peking University Press, 2002: 15-16.

Author Contributions

Su Ruizhu: Proposed the overall research ideas and framework; drafted and revised the paper; revised the final version. Yan Jingya: Responsible for data acquisition, collation, and analysis; drafted and revised the paper. Ouyang Jian: Responsible for paper revision; technical support for data processing.

Figures

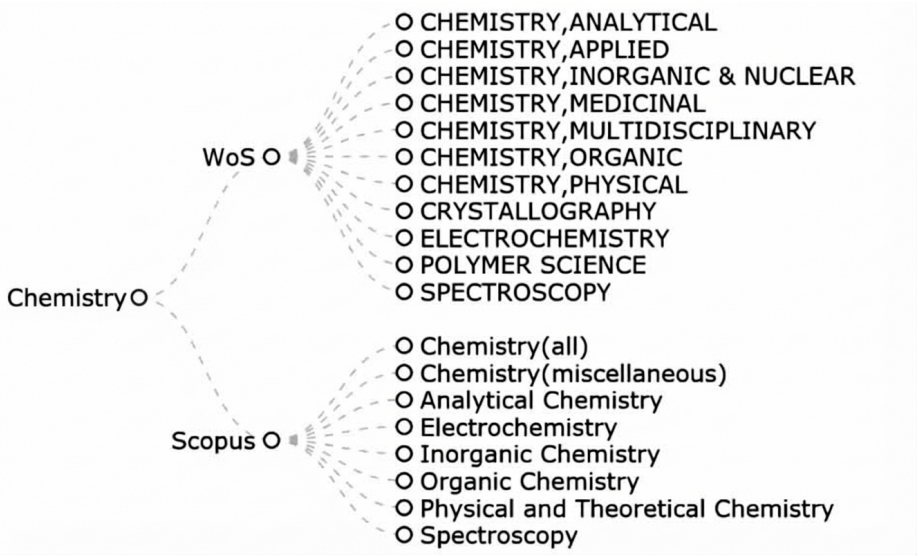

Figure 8: Figure 3

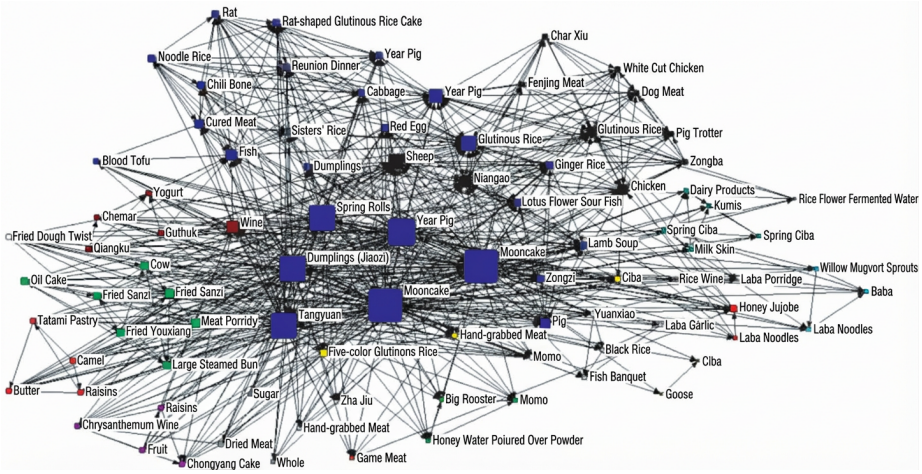

Figure 9: Figure 9